



## Course guide

### 295123 - 295II333 - Bioinformatics

Last modified: 08/08/2024

**Unit in charge:** Barcelona East School of Engineering  
**Teaching unit:** 707 - ESAII - Department of Automatic Control.

**Degree:** MASTER'S DEGREE IN INTERDISCIPLINARY AND INNOVATIVE ENGINEERING (Syllabus 2019). (Optional subject).  
ERASMUS MUNDUS MASTER'S DEGREE IN ADVANCED MATERIALS SCIENCE AND ENGINEERING (Syllabus 2021). (Optional subject).

**Academic year:** 2024    **ECTS Credits:** 6.0    **Languages:** English

#### LECTURER

---

**Coordinating lecturer:** ALEXANDRE PERERA LLUNA

**Others:**

#### PRIOR SKILLS

---

Basic Programming Skills

#### DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

---

**Specific:**

CEMUEII-18. Design innovative solutions in biomedicine through the use of computer tools for design, modeling and computational simulation (Specific competence of the Healthcare and Biomedical Applications specialty).

**Generical:**

CGMUEII-01. Participate in technological innovation projects in multidisciplinary problems, applying mathematical, analytical, scientific, instrumental, technological and management knowledge.

CGMUEII-05. To communicate hypotheses, procedures and results to specialized and non-specialized audiences in a clear and unambiguous way, both orally and through reports and diagrams, in the context of the development of technical solutions for problems of an interdisciplinary nature.

**Transversal:**

05 TEQ. TEAMWORK. Being able to work as a team player, either as a member or as a leader. Contributing to projects pragmatically and responsibly, by reaching commitments in accordance to the resources that are available.

06 URI. EFFECTIVE USE OF INFORMATION RESOURCES. Managing the acquisition, structure, analysis and display of information from the own field of specialization. Taking a critical stance with regard to the results obtained.

03 TLG. THIRD LANGUAGE. Learning a third language, preferably English, to a degree of oral and written fluency that fits in with the future needs of the graduates of each course.

#### TEACHING METHODOLOGY

---

The methodology of the course combines theory lessons, laboratory sessions and autonomous learning through the development of projects and the analysis of real applications.

#### LEARNING OBJECTIVES OF THE SUBJECT

---

-Properly design a bioinformatics analysis workflow, paying attention to the specific issues of the input data type to be analyzed



## STUDY LOAD

Type	Hours	Percentage
Hours large group	21,0	14.00
Self study	108,0	72.00
Hours small group	21,0	14.00

Total learning time: 150 h

## CONTENTS

### Biology background

**Description:**

- Basics of genetics
- Genetic variability and genetic diseases
- Genomic data and genomic landscape
- Genome databases and browsers

**Specific objectives:**

Learn key aspects of genetics: structure and function of nucleic acids and chromosomes.

- Learn the influence of genetics on diseases
- Know current types of functional genomic data: from 1D to 3D genome perspectives.
- Understand basic annotations in a genome. GENCODE project.
- Understand and interpret the information retrieved from a genome browser.

**Related activities:**

- Home activity 1: Analysis of mRNA sequences using command-line
- Laboratory session 1: Genome databases and browsers

**Full-or-part-time:** 18h

Theory classes: 4h

Laboratory classes: 2h

Self study : 12h

### Sequence alignment

**Description:**

- Local and global alignment algorithms
- Database similarity searches
- Protein motifs

**Related activities:**

- Understand the implications of carrying out a local or global alignment.
- Know the family of BLAST algorithms, local alignment tools.
- Get familiar with substitution matrices: PAM and BLOSUM families
- Understand the steps of progressive alignment.
- Identification of conserved motifs given several sequences.

**Full-or-part-time:** 26h

Theory classes: 4h

Laboratory classes: 6h

Self study : 16h



### Analysis of Transcriptomics Data

**Description:**

- Gene expression and regulation. Techniques and tools
- Processing: normalization, filtering and clustering
- Gene expression databases

**Related activities:**

- Laboratory 4: microarrays
- Laboratory 5: RNA-sequencing

**Full-or-part-time:** 26h

Theory classes: 4h

Laboratory classes: 6h

Self study : 16h

### Knowledge databases and enrichment analysis

**Description:**

- Biological pathways: KEGG, Reactome
- Gene ontologies: GO
- Protein interactions: STRINGdb
- Enrichment analysis

**Related activities:**

- Laboratory 6: Biological pathways and protein interactions
- Laboratory 7: Gene Ontology. Enrichment analysis

**Full-or-part-time:** 22h

Theory classes: 4h

Laboratory classes: 6h

Self study : 12h

### Advanced techniques and applications

**Description:**

- Proteomics and metabolomics
- Epigenomics
- Machine learning methods

**Related activities:**

- Laboratory 8: metabolomics analysis pipeline
- Home assignment 2: epigenomics

**Full-or-part-time:** 20h

Theory classes: 2h

Laboratory classes: 2h

Self study : 16h

## GRADING SYSTEM

Laboratory reports: 40%

Homework assignments: 20%

Final application project: 40%



## BIBLIOGRAPHY

---

### Basic:

- Compeau, Phillip; Pevzner, Pavel. Bioinformatics algorithms : an active learning approach (Vol. 1). 2nd ed. La Jolla, CA: Active Learning Publishers, 2015. ISBN 9780990374619.
- Compeau, Phillip; Pevzner, Pavel. Bioinformatics algorithms : an active learning approach (Vol. 2). 2nd ed. La Jolla, CA: Active Learning Publishers, 2015. ISBN 9780990374626.

## RESOURCES

---

### Hyperlink:

- KEGG. KEGG: Kyoto encyclopedia of genes and genomes
- Reactome. Reactome
- Bioconductor. Bioconductor
- Gene Ontology. Gene Ontology